

Conditional analysis for mixed covariates, with application to feed intake of lactating sows

SY Park¹, C Li¹, Z Han¹, SM Mendoza², E van Heugten², and AM Staicu¹

¹Department of Statistics, North Carolina State University

²Department of Animal Science, North Carolina State University

email: spark13@ncsu.edu

Friday 20th May, 2016

Abstract: We propose a novel modeling framework to study the effect of covariates of various types on the conditional distribution of the response. The methodology accommodates flexible model structure, allows for joint estimation of the quantiles at all levels, and involves a computationally efficient estimation algorithm. Extensive numerical investigation confirms good performance of the proposed method. The methodology is motivated by and applied to a lactating sow study, where the primary interest is to understand how the dynamic change of minute-by-minute temperature in the farrowing rooms within a day (functional covariate) is associated with low quantiles of feed intake of lactating sows, while accounting for other sow-specific information (vector covariate).

Keywords and phrases: Lactating sows, Functional Data Analysis, Penalized Regression, Partial Functional Quantile Regression..

1. Introduction

Many modern applications routinely collect data on study participants comprising scalar responses and both [vector and data streams] information and the main question of interest is to examine how the covariates affect the response. For example in our data application the aim is to study how the daily temperature or humidity behavior in the farrowing rooms (rooms where piglets are born and nursed by the sow until they are weaned) affect the feed intake of lactating sows during their first 21 lactation days, while accounting for other sow-specific information. A popular approach in these cases is to use a semi-parametric framework and assume that the mixed covariates solely affect the mean response; see Cardot et al. (1999); Ramsay and Silverman (2002); James (2002); Ferraty and Vieu (2006, 2009); Goldsmith et al. (2012); McLean et al. (2014) and others.

While it is still important to study the average feed intake, animal scientists are often

more concerned with the left tail of the feed intake distribution because low feed intake of lactating sows could lead to many serious issues, including decrease in milk production and negative impact on the sows reproductive system; see, for reference, Quiniou and Noblet (1999); Renaudeau and Noblet (2001); St-Pierre et al. (2003) among others. In this paper we relax the mean dependence assumption and consider that the covariates affect the entire conditional distribution of the response. Our primary objective is to develop a modeling framework for a comprehensive study of mixed covariates on scalar response; the sow data application includes daily measured temperature and humidity recorded at five minute intervals (functional type) as well as sow-specific information (vector type).

Quantile regression models the effect of scalar/vector covariates beyond the mean response and has attracted great interest (Koenker and Bassett Jr, 1978; Koenker, 2005). This approach offers a more comprehensive picture of the effects of the covariates on the response distribution. For pre-specified quantile levels, quantile regression models the conditional quantiles of the response as a function of the observed covariates; these approaches have been extended more recently to ensure non-crossing of quantile functions (Bondell et al., 2010). Quantile regression has been also extended to handle functional covariates. Cardot et al. (2005) discussed quantile regression models by employing a smoothing spline modeling-based approach. Kato et al. (2012) considered the same problem and used a functional principal component (fPC) based approach. Both papers mainly discussed the case of having a single functional covariate and it is not clear how to extend them to the case where there are multiple functional covariates or mixed covariates (vector and functional). Ferraty et al. (2005) and Chen and Müller (2012) considered a different perspective and studied the effect of a functional predictor on the quantiles of the response by first positing a model for the conditional distribution of the response and then inverting it; this approach is appealing as it jointly estimate quantiles for all the desired levels.

More recently, Tang and Cheng (2014), Lu et al. (2014), and Yu et al. (2015) studied quantile regression when the covariates are of mixed types and introduced the partial

functional linear quantile regression model framework. The first two papers used fPC basis while the last one proposed to use partial quantile regression (PQR) basis. While such approach is helpful when we are interested in studying the effect of covariate at particular quantile level, it provides an incomplete picture if we are interested in the effect at several quantile levels due to the well-known crossing-issue.

In this paper we propose a comprehensive study of the effect of vector and functional covariates on the distribution of the response. Our approach is inspired from Chen and Müller (2012) (CM, henceforth). Specifically let $Q_{Y|X}(\tau|X)$ denote the τ th conditional quantile of Y given a functional covariate $X(\cdot)$, and let $F_{Y|X}(y)$ denote the conditional distribution of Y given X . Using the relationship between $Q_{Y|X}(\tau|X)$ and $F_{Y|X}(y)$ that $Q_{Y|X}(\tau|X) = \inf\{y : F_{Y|X}(y) \geq \tau\}$ for $0 < \tau < 1$, CM proposed to estimate the quantile function $Q_{Y|X}(\tau|X)$ in two steps: 1) estimate the conditional distribution of Y given $X(\cdot)$, $F(y) = E[\mathbb{1}(Y < y)|X(\cdot)]$ by positing a mean regression model for an auxiliary variable $Z(y) := \mathbb{1}(Y \leq y)$ and the functional covariate $X(\cdot)$; and 2) estimate $Q_{Y|X}(\tau|X)$ by inverting the estimated conditional distribution function. Their estimation approach is restrictive to one functional covariate and a direct extension to accommodate mixed covariates is computationally expensive. We consider a similar idea and propose a modeling framework and estimation technique that easily accommodate various types of covariates in a computationally efficient manner. This development represents the main contribution of our manuscript. Let X_1 be a scalar covariate and $X_2(\cdot)$ be a functional covariate defined on a closed domain \mathcal{T} . We propose the following model for the conditional distribution of Y given X_1 and $X_2(\cdot)$:

$$E[\mathbb{1}(Y < y)] = g^{-1}\{\beta_0(y) + X_1\beta_1(y) + \int X_2(t)\beta_2(t, y)dt\}, \quad (1)$$

where $g(\cdot)$ is a known, monotone link function, $\beta_1(y)$ is unknown and smooth function and $\beta_2(t, y)$ is unknown and smooth bi-variate function over y and t . The parameters $\beta_1(y)$ and $\beta_2(\cdot, y)$ quantify the effect of the covariates X_1 and $X_2(\cdot)$ respectively onto the

distribution of the response.

The remainder of the paper is structured as follows. Section 2 discusses the details of the proposed method and Section 3 performs a thorough simulation study evaluating the performance of the proposed method and competitors. We apply the proposed method to analyze the sow data in Section 4. We conclude the paper with a discussion in Section 5.

2. Modeling Framework

The proposed modeling and estimation method is discussed first for the case of a scalar covariate in Section 2.1; Section 2.2 considers the extension to the case of a functional covariate and then to the case of mixed covariates; Section 2.3 further extends the method to handle sparse and noisy functional covariates. We briefly discuss the monotonization of the estimated conditional distribution in Section 2.4.

2.1. Conditional distribution of the response given scalar covariate

First, we focus on the case of a scalar covariate X . Consider the data $\{(X_i, Y_i) : i = 1, \dots, n\}$, where X_i and Y_i are independent realizations of real-valued scalar random variables X and Y , respectively. Define $Z_i(y) = \mathbb{1}(Y_i < y)$ for $y \in \mathbb{R}$, where $\mathbb{1}(\cdot)$ is an indicator function; for each y , we view $Z_i(y)$ as a binary-valued random variable that is independent and identically distributed as $Z(y) = \mathbb{1}(Y < y)$. It follows that the conditional distribution function $F_{Y|X}(y) = \mathbb{E}[Z(y)|X]$. Here we propose to model the conditional distribution, $F_{Y|X}(y)$, using a generalized function-on-scalar regression model (Goldsmith et al., 2011) between the ‘artificial’ functional response $Z_i(y)$ and the scalar covariate X_i . Specifically, for each $y \in \mathbb{R}$, consider

$$\mathbb{E}[Z(y)|X] = g^{-1}\{\beta_0(y) + X\beta_1(y)\}, \quad (2)$$

where $g(\cdot)$ is a known, monotonic link function, and $\beta_0(\cdot)$ and $\beta_1(\cdot)$ are unknown, smooth coefficient functions. Here we use the logit function defined as $g(x) = \log\{x/(1+x)\}$. It

is noteworthy to remark that if the slope parameter $\beta_1(\cdot)$ is null then the covariate X has no effect on the distribution of the response Y , which is equivalent to X having no effect on any quantile level of Y .

We model $\beta_0(y)$ and $\beta_1(y)$ by using pre-specified, truncated univariate basis, $\{B_{0,d}(\cdot) : d = 1, \dots, \kappa_0\}$ and $\{B_{1,d}(\cdot) : d = 1, \dots, \kappa_1\}$: $\beta_0(y) = \sum_{d=1}^{\kappa_0} B_{0,d}(y)\theta_{0,d}$ and $\beta_1(y) = \sum_{d=1}^{\kappa_1} B_{1,d}(y)\theta_{1,d}$, where $\theta_{0,d}$'s and $\theta_{1,d}$'s are unknown basis coefficients. Then model (2) can be represented as the following generalized additive model

$$E[Z(y)|X] = g^{-1}\left\{\sum_{d=1}^{\kappa_0} B_{0,d}(y)\theta_{0,d} + \sum_{d=1}^{\kappa_1} B_{1,d}(y)\theta_{1,d}\right\}, \quad (3)$$

where for convenience we use the notation $B_{x,d}(y) = XB_{1,d}(y)$. The general idea is to set the basis dimensions κ_0 and κ_1 to be sufficiently large to capture the complexity of the coefficient functions and control the smoothness of the estimator through roughness penalties $P_0(\boldsymbol{\theta}_0)$ and $P_1(\boldsymbol{\theta}_1)$, where $\boldsymbol{\theta}_l$, is a vector of all basis coefficients $\{\theta_{l,d} : l = 1, \dots, D_l\}$ for $l = 1, 2$. This approach of using roughness penalties has been widely used; see, for example, Eilers and Marx (1996); Ruppert (2002); Wood (2003, 2006a) among many others.

In the following, we detail the estimation algorithm. Let $\{y_j : j = 1, \dots, J\}$ be a set of equi-spaced points in the range of the response variable, Y_i 's. For each i and j , we define $Z_{ij} = Z_i(y_j) = \mathbb{1}(Y_i < y_j)$; it follows that conditional on X_i , the Z_{ij} are independently distributed as Bernoulli distribution with mean (μ_{ij}) , where μ_{ij} is such that $g(\mu_{ij}) = \mathbf{B}_{0,j}^T \boldsymbol{\theta}_0 + \mathbf{B}_{x,j}^T \boldsymbol{\theta}_1$. Here $\mathbf{B}_{0,j}^T$ is a $\kappa_0 \times 1$ vector of $\{B_{0,d}(y_j) : d = 1, \dots, \kappa_0\}$ and $\mathbf{B}_{x,j}^T$ is a $\kappa_1 \times 1$ vector of $\{X_i B_{1,d}(y_j) : d = 1, \dots, \kappa_1\}$. The basis coefficients, $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$, are estimated by maximizing the penalized log likelihood criterion,

$$2\log\mathcal{L}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1 | \{Z_{ij}(y_j), X_i : \forall i, j\}) - \lambda_0 P_0(\boldsymbol{\theta}_0) - \lambda_1 P_1(\boldsymbol{\theta}_1), \quad (4)$$

where \mathcal{L} is the likelihood function of data $\{Z_{ij} : j = 1, \dots, J\}_i$, $P_0(\boldsymbol{\theta}_0)$ and $P_1(\boldsymbol{\theta}_1)$ are penalties, and λ_0 and λ_1 are smoothing parameters. We use quadratic penalties which

penalize the size of the curvature of the estimated coefficient functions. Let $P_0(\boldsymbol{\theta}_0) = \boldsymbol{\theta}_0^T \mathbf{D}_0 \boldsymbol{\theta}_0$ and $P_1(\boldsymbol{\theta}_1) = \boldsymbol{\theta}_1^T \mathbf{D}_1 \boldsymbol{\theta}_1$, where \mathbf{D}_0 and \mathbf{D}_1 are $\kappa_0 \times \kappa_0$ and $\kappa_1 \times \kappa_1$ dimensional matrices based on the basis used (see Wood (2006a) for example; the (s, s') element of \mathbf{D}_0 is $\int B''_{0,s}(y)B''_{0,s'}(y)dy$) and \mathbf{D}_1 is defined similarly. The smoothing parameters λ_0 and λ_1 control the trade-off between the goodness of fit and smoothness of the fit. The smoothing parameters are estimated using restricted maximum likelihood (REML).

The criterion (4) can be viewed as the penalized quasi-likelihood (PQL) of the corresponding generalized linear mixed model

$$Z_{ij}|\boldsymbol{\theta}_0, \boldsymbol{\theta}_1 \sim \text{Bernoulli}(\mu_{ij}); \quad \boldsymbol{\theta}_0 \sim N(\mathbf{0}, \lambda_0^{-1} \mathbf{D}_0^-); \quad \boldsymbol{\theta}_1 \sim N(\mathbf{0}, \lambda_1^{-1} \mathbf{D}_1^-), \quad (5)$$

where \mathbf{D}_0^- is the generalized inverse matrix of \mathbf{D}_0 and \mathbf{D}_1^- is defined similarly. Wood (2006a) discusses an alternative way to deal with the rank-deficient matrices, \mathbf{D}_0 and \mathbf{D}_1 , in the context of restricted maximum likelihood (REML) estimation. See also Ivanescu et al. (2015) who uses the mixed model representation of a similar regression model to (5), but with a Gaussian functional response.

Let $\{\hat{\theta}_{l,d} : d = 1, \dots, \kappa_l\}$ for $l = 0, 1$ be the estimated basis coefficients. It follows that the estimated distribution function is $\hat{F}_{Y|X}(y) = g^{-1}\{\sum_{d=1}^{\kappa_0} B_{0,d}(y)\hat{\theta}_{0,d} + \sum_{d=1}^{\kappa_1} B_{x,d}(y)\hat{\theta}_{1,d}\}$. The τ th conditional quantile are estimated as by inverting the estimated distribution $\hat{F}_{Y|X}(\cdot)$, $\hat{Q}_{Y|X}(\tau|X) = \min_j \{y_j : \hat{F}_{Y|X}(y_j) \geq \tau\}$. This approach relates the τ th level quantile of the response in a nonlinear manner to the covariate. One should note that the estimated distribution function, $\hat{F}_{Y|X}(y)$, here is not a monotonic function yet. However in practice one can always obtain $\hat{F}_{Y|X}(y)$ using a monotonization method as described in Section 2.4 first and then invert the resulting estimated distribution to get the estimated conditional quantiles.

2.2. Extension to mixed covariates

The modeling approach discussed in Section 2.1 is quite powerful as it can accommodate covariates of various types, including functional covariates.

Assume that a functional covariate $X_i(\cdot)$ is observed. For convenience, we assume that $X_i(\cdot)$ is observed at fine grid of points and without noise after all. Instead of generalized function-on-scalar model as is used in Section 2.1, now we use generalized function-on-function regression model $E[Z(y)|X] = g^{-1}\{\beta_0(y) + \int X(t)\beta_1(t, y)dt\}$, where $\beta_1(\cdot, \cdot)$ is an unknown bi-variate coefficient function. In terms of modeling and estimation, the main difference from Section 2.1 is that the coefficient function, $\beta_1(t, y)$, is now bivariate and it requires appropriate pre-specified basis function and corresponding penalty term.

We represent $\beta_1(t, y)$ using the tensor product of two univariate bases functions, $\{B_{1,d_t}^t(t) : d_t = 1, \dots, \kappa_{1,t}\}$ and $\{B_{1,d_y}^y(y) : d_y = 1, \dots, \kappa_{1,y}\}$; $\beta_1(t, y) = \sum_{d_t=1}^{\kappa_{1,t}} \sum_{d_y=1}^{\kappa_{1,y}} B_{1,d_t}^t(t) B_{1,d_y}^y(y) \theta_{1,d_t,d_y}$. Subsequently the previous penalty matrix \mathbf{D}_1 should be also appropriately modified to control the smoothness of $\beta_1(t, y)$ in directions of both t and y . There are several choices to define the penalty matrix in nonparametric regression (see Marx and Eilers (2005); Xiao et al. (2013)). For bivariate smoothing we use $\mathbf{D}_1 = \{\mathbf{P}_t \otimes \mathbf{I}_{\kappa_{1,y}} + \mathbf{I}_{\kappa_{1,t}} \otimes \mathbf{P}_y\}$, where (s, s') element of \mathbf{P}_t is $\int \{\partial^2 X(t) B_{1,s}^t(t) / \partial t^2\} \{\partial^2 X(t) B_{1,s'}^t(t) / \partial t^2\} dt$ and (s, s') element of \mathbf{P}_y is $\int \{\partial^2 B_{1,s}^y(y) / \partial y^2\} \{\partial^2 B_{1,s'}^y(y) / \partial y^2\} dy$ introduced by Wood (2006a,b). In practice the integration term $\int X(t)\beta_1(t, y)dt$ is approximated by Riemann integration $\int X(t)\beta_1(t, y)dt = \sum_{l=1}^L X(t_l)\beta_1(t_l, y)(t_{l+1} - t_l)$, but other numerical approximation can be also used. The estimation of parameters proceeds similarly to Section 2.1.

The ideas can be further extended to accommodate multiple covariates, scalar or functional, and varied types of effects. For example, assume that we have one vector covariate, one scalar covariate and one functional covariate. We posit a model that considers that the vector covariate has constant effect on the response while both the scalar and the

functional covariates have varying effect on the response. Consider the model:

$$F_{Y|X}(y) = g^{-1}\{\beta_0(y) + \mathbf{X}_1^T \boldsymbol{\beta}_1 + X_2 \beta_2(y) + \int X_3(t) \beta_3(t, y) dt\}, \quad (6)$$

for covariate vector \mathbf{X}_1 , scalar covariate X_2 , and functional covariate $X_3(\cdot)$. Here \mathbf{X}_1 is assumed to have the y -invariant linear effect, and X_2 and X_3 are assumed to have the y -variant linear effects. It is easy to see that a null effect, say $\beta_3(\cdot, \cdot) \equiv 0$ is equivalent to the fact that the corresponding covariate, in this case $X_3(\cdot)$ has no effect on any quantile levels of the response.

Fitting models given in (2) and (6) can be done by extending the ideas of Ivanescu et al. (2015), which provide general modeling and estimation methods for penalized function-on-function regression for the case of Gaussian functional response and implement their method in R (R Core Team, 2013) (namely, the `pffr` function in `refund` package (Crainiceanu et al., 2013)). The extension of the model to the non-Gaussian response has recently been studied and implemented by Scheipl et al. (2015). Based on the existing function we implement the proposed method and provide a wrapper function in R.

Furthermore the proposed method can be easily extended to relax the linearity assumption and allow more flexible model structures. For example, instead of a functional linear model such as $F_{Y|X}(y) = g^{-1}\{\beta_0(y) + X_1 \beta_1(y) + \int X_2(t) \beta_2(t, y) dt\}$, we can model the conditional distribution as $F_{Y|X}(y) = g^{-1}\{\beta_0(y) + h_1(X_1) + \int h_2(X_2(t), t, y) dt\}$, where $h_1(\cdot)$ and $h_2(\cdot, \cdot, \cdot)$ are unknown univariate and trivariate smooth functions, respectively. We illustrate the nonlinear model in the simulation study for the case when there is a single scalar covariate and the corresponding results are presented in Section S1.1 of the Supplementary Materials. The results show excellent prediction performance as the competitive nonlinear quantile regression method, namely Constrained B-Spline Smoothing (Ng and Maechler, 2007), denoted by COBS.

It is important to emphasize that even in the case of a single functional covariate, our methodology differs from Chen and Müller (2012) (CM) in few directions: 1) Our proposed

method is based on modeling the unknown smooth coefficient functions using pre-specified basis function expansion and using penalties to control their roughness. In contrast, CM uses data-driven basis, chooses the number of basis functions through the percentage of explained variance (PVE) of the functional predictors. This key difference allows us to accommodate covariates of different types. 2) Our estimation approach is based on a single step penalized criterion while CM uses pointwise estimation based on the residual sum of square criterion and thus requires fitting multiple generalized regressions. This is an important advantage in terms of computational efficiency.

2.3. Extension to sparse and noisy functional covariates

In practice the functional covariates are often observed at irregular times across the units and also are possibly corrupted with measurement errors. In such case, one needs to first smooth and de-noise the trajectories before fitting. When the sampling design of the functional covariate is dense, then the common approach is to take each trajectory and smooth it using spline or local polynomial smoothing, as proposed in Ramsay (2006) and Zhang et al. (2007). When the design is sparse, the smoothing is done by pooling all the subjects and using the method proposed in Yao et al. (2005). A method following Yao et al. (2005) (`fpca.sc`) is implemented in R package `refund` (Crainiceanu et al., 2013). In our simulation study we use `fpca.sc` for this step irrespective of a sampling design (dense or sparse); alternatively, `fpca.face` (Xiao et al., 2016) in `refund` or `PACE` (Yao et al., 2005) in `MATLAB` (MATLAB, 2010) can also be used.

2.4. Monotonization

While a conditional quantile function is nondecreasing, the resulting estimated quantiles may not be. We consider monotonization as proposed by Chernozhukov et al. (2009). Chernozhukov et al. (2009) showed that in this way the monotonized estimator gives the same or better fit than the original estimator. Two approaches are widely used; one is to

monotonize the estimated conditional distribution function $\widehat{F}_{Y|X}(y)$, and the other is to monotonize the estimated conditional quantile function $\widehat{Q}_{Y|X}(\tau)$. We choose the former as we already have $\widehat{F}_{Y|X}(y)$ evaluated at dense grid points y_j 's and there is no need to obtain the estimated conditional quantile at fine grid points of the quantile level $\tau \in [0, 1]$. We use an isotonic regression model (Barlow et al., 1972) for monotonization as it makes no structural assumption and gives an ordered fit; It fits a nonparametric model with an order restriction. The isotonic regression model is fitted through $\{(y_j, \widehat{F}_{Y|X}(y_j)) : j = 1, \dots, J\}$ using the `isoreg` function in R (R Core Team, 2013). This idea was also employed in Kato et al. (2012).

3. Simulation study

In this section we evaluate numerically the performance of the proposed method. We present results for the case when we have both functional and scalar covariates; additional results when there is only single functional or single scalar covariate are discussed in the Supplementary Materials, Section S1. We adapt the simulation settings of Chen and Müller (2012) for the cases that involve a functional covariate.

Suppose the observed data for the i th subject is $[Y_i, X_{1i}, \{(W_{i1}, t_{i1}), \dots, (W_{im_i}, t_{im_i})\}]$, $t_{ij} \in [0, 10]$, where $X_{1i} \stackrel{i.i.d}{\sim} Unif(-16, 16)$, $W_{ij} = X_{2i}(t_{ij}) + \epsilon_{ij} = \mu(t_{ij}) + \sum_{k=1}^4 \xi_{ik} \phi_k(t_{ij}) + \epsilon_{ij}$, $1 \leq i \leq n$, $1 \leq j \leq m_i$. Set the mean function $\mu(t) = t + \sin(t)$, and the eigenfunctions $\phi_k(t) = \cos\{(k+1)\pi t/10\}/\sqrt{5}$ for odd values of k , $\phi_k(t) = \sin\{k\pi t/10\}/\sqrt{5}$ for even values of k . Here, assume that scores $\xi_{ik} \stackrel{iid}{\sim} N(0, \lambda_k)$, where $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \{16, 9, 7.56, 5.06\}$, and $\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$. We assume two cases:

- (i) normal distribution $Y_i|X_{1i}, X_{2i}(\cdot) \sim N(2 \int X_{2i}(t)\beta(t)dt + 2X_{1i}, 5^2)$; this yields the quantile regression model $Q_{Y|X_1, X_2(\cdot)}(\tau) = 2 \int X_{2i}(t)\beta(t)dt + 2X_{1i} + 5\Phi^{-1}(\tau)$, where $\Phi(\cdot)$ is the distribution function of the standard normal and $\beta(t) = \sum_{k=1}^4 \beta_k \phi_k(t)$;
- (ii) mixture of normal distributions $Y_i|X_{1i}, X_{2i}(\cdot) \sim 0.5N(\int X_{2i}(t)\beta(t)dt + X_{1i}, 1^2) + 0.5N(3 \int X_{2i}(t)\beta(t)dt + 3X_{1i}, 4^2)$; this yields the quantile regression model $Q_{Y|X_1, X_2(\cdot)}(\tau) =$

$$2 \int X_{2i}(t)\beta(t)dt + 2X_{1i} + \sqrt{(\int X_{2i}(t)\beta(t)dt + X_{1i})^2 + 8.5\Phi^{-1}(\tau)}, \text{ where } \beta(t) = \sum_{k=1}^4 \beta_k \phi_k(t).$$

Three noise levels are considered: low ($\sigma_\epsilon = 0.50$), moderate ($\sigma_\epsilon = 4.33$), and high ($\sigma_\epsilon = 6.13$). The three levels are such that the signal to noise ratio (SNR), which are calculated as $SNR = \sqrt{\sum_{k=1}^4 \lambda_k} / \sigma_\epsilon$, are equal to $SNR = 150$, 2, and 1, respectively. Results are presented for sample sizes, $n = 100$ (small) and $n = 1000$ (large).

The performance is evaluated on a test set of size 100. Two sampling designs are considered: (i) *dense design*, where the sampling points $\{t_{ij} : j = 1, \dots, m_i\}$ are a set of $m_i = 30$ equi-spaced time points in $[0, 10]$; and (ii) *sparse design*, where $\{t_{ij} : j = 1, \dots, m_i\}$ are $m_i = 15$ randomly selected points from a set of 30 equi-spaced grids in $[0, 10]$. The quantile functions with our approach are estimated as described in Section 2 by first creating an artificial binary response $Z_i(\cdot)$ and then fitting a penalized function-on-function regression model and using the logit link function; we use the `pfr` function (Ivanescu et al., 2015; Scheipl et al., 2015) in the `refund` package (Crainiceanu et al., 2013) in R (R Core Team, 2013) for binomial responses, denoted by Joint QR.

We compare our method with three alternative approaches: (1) a variant of our proposed approach using pointwise fitting, denoted by Pointwise QR, and hence fitting multiple regression models with binomial link function as implemented by the penalized functional regression `pfr` of the `refund` package for generalized scalar responses, developed by Goldsmith et al. (2011); (2) a modified version of the CM method, denoted by Mod CM, that we developed to account for additional scalar covariates, and which estimates pointwise using multiple generalized linear models; (3) a linear quantile regression approach using the quantile loss function and the partial quantile regression bases for functional covariates, proposed by Yu et al. (2015) and denoted by PQR. Notice that (1) and (2) account for a varying effect of the covariates on the response distribution, but do not ensure that this effect is smooth.

The R function `pfr` can incorporate both scalar/vector and functional predictors by

adopting a mixed effects model framework. The functional covariates are pre-smoothed by fPC analysis and truncation is done based on the number of fPCs determined by a percentage of explained variance equal to 99% for all estimation methods; pre-smoothing the functional covariates before fitting the regression model has been also considered by Goldsmith et al. (2011) and Ivanescu et al. (2015). The performance is evaluated in terms of mean absolute error (MAE) for quantile levels $\tau = 0.05, 0.1, 0.25$, and 0.5 .

Numerical results are provided in Tables 1, 2 and 3. Table 1 gives results for the two settings (normal and mixture) when the functional covariate is observed on dense design and the sample size is $n = 100$; Table 2 shows the corresponding results for $n = 1000$. We obtained similar findings for the sparse scenario and hence are not reported for brevity.

Consider first the case when conditional distribution of the response is normal. When sample size is large ($n = 1000$) the proposed method (Joint QR) yields the best MAE for the SNR and the quantile levels considered. Even with low-moderate sample size ($n = 100$) the Joint QR remains performing the best for extreme quantiles ($\tau = 0.05, 0.10$ and 0.25) and relatively large noises ($\sigma_\epsilon = 4.33$ and 6.13). When sample size is small-moderate the PQR method also performs very well for the small noise ($\sigma_\epsilon = 0.50$) and for middle quantiles ($\tau = 0.05$ or $\tau = 0.10$). The Pointwise QR and Mod CM methods perform similarly, where the Pointwise QR tends to do better for low-moderate sample sizes ($n = 100$) while the Mod CM tends to do better for larger one ($n = 1000$). All of the four methods are affected by the level of SNR; the higher it is, the better MAE is.

When the conditional distribution of the response follows the mixture of normals, there is no uniformly best method across quantiles levels or SNR levels whe sample size is large ($n = 1000$). It seems that all four methods have similar performance with some being the best for some situations while others for other situations. Overall the Joint QR method tends to perform better for extreme quantiles ($\tau = 0.05$ or $\tau = 0.10$) while the other three methods tend to predict better the middle quantiles ($\tau = 0.25$ or $\tau = 0.50$). Other findings are relatively similar to the ones for the normal case.

Table 3 compares the three methods that involve estimating the conditional distribution - Joint QR, Pointwise QR and Mod CM - in terms of the computational time required for fitting; the times correspond to using a desktop computer with a 2.3 GHz CPU and 8 GB of RAM. Not surprisingly by fitting the model a single time, Joint QR is the fastest, in some cases being order of magnitude faster than the rest. Pointwise QR can be up to twice as fast as Mod CM.

For completeness, we also compare our proposed method to the appropriate competitive methods for the cases (1) when there is a single scalar covariate and (2) when there is a single functional covariate. The Supplementary Materials, Section S1.1 discusses the former case and compares Joint QR and Pointwise QR with the linear quantile regression and the nonlinear quantile regression (as implemented by the `cobs` function in the R package COBS (Ng and Maechler, 2007)) in an extensive simulation experiment that involves both linear quantile settings and nonlinear quantile settings. Overall the results show that the proposed methods have similar behavior as LQR; see Table S1. Furthermore we consider the proposed methods with nonlinear modeling of the conditional distribution as discussed in Section 2, which we denote with Joint QR (NL) for joint fitting and Pointwise QR (NL) for pointwise fitting. Nonlinear versions of the proposed methods have an excellent MAE performance, which is comparable to or better than that of the COBS method.

Finally, Section S1.2 in the Supplementary Materials discusses the simulation study for the case of having a single functional covariate and compares the proposed methods with CM in terms of MAE as well as computational time; see results displayed in Tables S2 and S3. The results show that the proposed Joint QR is superior to CM both in terms of the prediction accuracy and computation efficiency. In our simulation study we also consider the joint fitting of the model by treating the binary response as normal and use `pffr` (Ivanescu et al., 2015) with Gaussian link, denoted by Joint QR (G).

TABLE 1

Simulation results: Average MAE of the predicted τ -level quantile for sample size $n = 100$ (standard error in parentheses). Results are based on 500 replications.

| Distribution | σ_ϵ | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ |
|--------------|-------------------|--------------|---------------------|---------------------|--------------------|--------------------|
| Normal | 0.50 | Joint QR | 3.30(0.03) | 3.15(0.03) | 2.90(0.02) | 2.71(0.02) |
| | | Pointwise QR | 4.41 (0.03) | 4.06 (0.03) | 3.67 (0.02) | 3.59 (0.02) |
| | | Mod CM | 4.32 (0.03) | 4.04 (0.03) | 3.67 (0.02) | 3.53 (0.02) |
| | | PQR | 2.97 (0.04) | 2.72 (0.04) | 2.50 (0.04) | 2.45 (0.04) |
| Normal | 4.33 | Joint QR | 8.13 (0.04) | 7.57 (0.03) | 7.13 (0.03) | 7.02 (0.03) |
| | | Pointwise QR | 9.37 (0.05) | 8.46 (0.04) | 7.63 (0.03) | 7.36 (0.03) |
| | | Mod CM | 8.88 (0.04) | 8.69 (0.04) | 8.52 (0.04) | 8.60 (0.04) |
| | | PQR | 8.76 (0.06) | 8.04 (0.05) | 7.13 (0.04) | 6.76 (0.03) |
| Normal | 6.13 | Joint QR | 9.90 (0.05) | 9.05 (0.04) | 8.32 (0.03) | 8.13 (0.03) |
| | | Pointwise QR | 11.07 (0.06) | 9.90 (0.05) | 8.78 (0.03) | 8.41 (0.03) |
| | | Mod CM | 10.27 (0.04) | 10.00 (0.04) | 9.79 (0.04) | 9.84 (0.04) |
| | | PQR | 10.67 (0.07) | 9.71 (0.06) | 8.40 (0.04) | 7.86 (0.03) |
| Mixture | 0.50 | Joint QR | 6.59 (0.06) | 5.77 (0.06) | 6.17 (0.06) | 4.45 (0.05) |
| | | Pointwise QR | 7.53 (0.07) | 6.13 (0.06) | 6.07 (0.06) | 4.64 (0.06) |
| | | Mod CM | 6.66 (0.06) | 5.94 (0.06) | 6.37 (0.06) | 4.95 (0.06) |
| | | PQR | 8.08 (0.07) | 7.06 (0.05) | 6.32 (0.04) | 5.91 (0.14) |
| Mixture | 4.33 | Joint QR | 10.40 (0.06) | 9.06 (0.05) | 8.62 (0.05) | 6.89 (0.05) |
| | | Pointwise QR | 11.70 (0.08) | 9.65 (0.06) | 8.61 (0.05) | 6.88 (0.05) |
| | | Mod CM | 11.40 (0.06) | 10.89 (0.06) | 10.68 (0.06) | 9.34 (0.07) |
| | | PQR | 12.00 (0.08) | 10.28 (0.06) | 8.68 (0.04) | 6.09 (0.06) |
| Mixture | 6.13 | Joint QR | 11.79 (0.07) | 10.15 (0.05) | 9.38 (0.05) | 7.53 (0.05) |
| | | Pointwise QR | 12.94 (0.09) | 10.68 (0.06) | 9.37 (0.05) | 7.48 (0.05) |
| | | Mod CM | 12.61 (0.07) | 11.95 (0.06) | 11.60 (0.06) | 10.23 (0.07) |
| | | PQR | 13.40 (0.10) | 11.47 (0.07) | 9.48 (0.04) | 6.59 (0.05) |

TABLE 2

Simulation results: Average MAE of the predicted τ -level quantile for sample size $n = 1000$ (standard error in parentheses). Results are based on 500 replications.

| Distribution | σ_ϵ | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ |
|--------------|-------------------|--------------|---------------------|--------------------|--------------------|--------------------|
| Normal | 0.50 | Joint QR | 1.27 (0.01) | 1.34 (0.01) | 1.33 (0.01) | 1.22 (0.01) |
| | | Pointwise QR | 1.61 (0.01) | 1.59 (0.01) | 1.54 (0.01) | 1.43 (0.01) |
| | | Mod CM | 1.43 (0.01) | 1.40 (0.01) | 1.39 (0.01) | 1.36 (0.01) |
| | | PQR | 1.74 (0.02) | 1.71 (0.02) | 1.67 (0.02) | 1.67 (0.02) |
| Normal | 4.33 | Joint QR | 7.79 (0.03) | 6.91 (0.02) | 6.22 (0.02) | 6.11 (0.02) |
| | | Pointwise QR | 8.03 (0.03) | 7.06 (0.03) | 6.31 (0.02) | 6.19 (0.02) |
| | | Mod CM | 7.94 (0.03) | 7.11 (0.03) | 6.43 (0.02) | 6.32 (0.02) |
| | | PQR | 8.36 (0.04) | 7.55 (0.04) | 6.60 (0.03) | 6.22 (0.02) |
| Normal | 6.13 | Joint QR | 9.84 (0.04) | 8.59 (0.03) | 7.52 (0.03) | 7.30 (0.03) |
| | | Pointwise QR | 10.09 (0.04) | 8.76 (0.03) | 7.61 (0.03) | 7.35 (0.03) |
| | | Mod CM | 9.95 (0.04) | 8.79 (0.03) | 7.77 (0.03) | 7.55 (0.03) |
| | | PQR | 10.39 (0.05) | 9.26 (0.04) | 7.89 (0.03) | 7.33 (0.02) |
| Mixture | 0.50 | Joint QR | 4.33 (0.03) | 2.89 (0.02) | 4.62 (0.03) | 3.54 (0.03) |
| | | Pointwise QR | 4.04 (0.03) | 2.42 (0.02) | 4.25 (0.03) | 3.45 (0.03) |
| | | Mod CM | 4.12 (0.03) | 2.28 (0.02) | 4.08 (0.03) | 3.65 (0.03) |
| | | PQR | 7.74 (0.04) | 6.40 (0.03) | 5.31 (0.02) | 3.75 (0.10) |
| Mixture | 4.33 | Joint QR | 9.62 (0.04) | 7.46 (0.03) | 7.10 (0.03) | 5.74 (0.03) |
| | | Pointwise QR | 9.64 (0.04) | 7.37 (0.03) | 6.86 (0.03) | 5.67 (0.03) |
| | | Mod CM | 9.64 (0.04) | 7.60 (0.03) | 7.18 (0.03) | 6.12 (0.03) |
| | | PQR | 11.74 (0.05) | 9.65 (0.04) | 7.82 (0.03) | 4.64 (0.02) |
| Mixture | 6.13 | Joint QR | 11.47 (0.04) | 8.94 (0.03) | 8.00 (0.03) | 6.33 (0.03) |
| | | Pointwise QR | 11.55 (0.05) | 8.91 (0.03) | 7.82 (0.03) | 6.22 (0.03) |
| | | Mod CM | 11.48 (0.04) | 9.17 (0.04) | 8.22 (0.04) | 6.81 (0.03) |
| | | PQR | 13.33 (0.06) | 10.87 (0.04) | 8.65 (0.03) | 5.37 (0.02) |

TABLE 3

Comparison of the average computing time (seconds) for the three approaches that involve estimating the conditional distribution.

| Distribution | Method | $n = 100$ | $n = 1000$ |
|--------------|--------------|-----------|------------|
| Normal | Joint QR | 17 | 200 |
| | Pointwise QR | 148 | 271 |
| | Mod CM | 278 | 511 |
| Mixture | Joint QR | 18 | 282 |
| | Pointwise QR | 151 | 296 |
| | Mod CM | 327 | 532 |

4. Sow Data Application

Our motivating application is an experimental study carried out at a commercial farm in Oklahoma from July 21, 2013 to August 19, 2013 (Rosero et al., 2016). The study comprises of 480 lactating sows of different parities (i.e. the number of previous pregnancies, which serves as a surrogate for age and body weight) that were observed during their first 21 lactation days; their feed intake was recorded daily as the difference between the feed offer and the feed refusal. In addition the study contains information on the temperature and humidity of the farrowing rooms, each recorded at five minute intervals. The final dataset we used for the analysis consists of 475 sows after five sows with unreliable measurements were removed by the experimenters.

The experiment was conducted to gain better insights into the way that the ambient temperature and humidity of the farrowing room affects the feed intake of lactating sows. Previous studies seem to suggest a reduction in the sow's feed intake due to heat stress: above 29°C sows decrease feed intake by 0.5 kg per additional degree in temperature Quiniou and Noblet (1999). Studying the effect of heat stress on lactating sows is a very important scientific question because of a couple of reasons. First, the reduction of feed intake of the lactating sows is associated with a decrease in both their bodyweight (BW) and milk production, as well as the weight gain of their litter (Johnston et al., 1999; Renaudeau and Noblet, 2001; Renaudeau et al., 2001). Sows with poor feed intake during lactation continue the subsequent reproductive period with negative energy balance (Black et al., 1993), which acts as negative feedback to prevent the onset of a new reproductive cycle. Second, heat stress reduced farrowing rate (the number of sows that deliver a new litter) and reduced the number of piglets born (Bloemhof et al., 2013); The reduction in reproductivity due to seasonality is estimated to cost 300 million dollars per year for the swine industry (St-Pierre et al., 2003). Economic losses are estimated to increase (Nelson et al., 2009) because very high temperatures are likely to occur more frequently due to

global warming (Melillo et al., 2014).

Our primary goal is to understand the thermal needs of the lactating sows for proper feeding behavior during the lactation time. Specifically, we are interested in how the interplay between the temperature and humidity of the farrowing room affects the feed intake demeanor of lactating sows of different parities. We focus on three specific times during the lactation period: beginning (lactation day 4), middle (day 11) and end (lactation day 18) and consider two types of responses that are meant to assess the feed intake behavior using the current and the previous lactation day. The first one quantifies the absolute change in the feed intake over two consecutive days and the second one quantifies the relative change and takes into account the usual sow's feed intake. We define them formally after introducing some notation.

Let FI_{ij} be the j th measurement of the feed intake observed for the i th sow and denote by LD_{ij} the lactation day when FI_{ij} is measured; here $j = 1, \dots, n_i$. Most sows are observed for every day within the first 21 lactation days and thus have $n_i = 21$. First define the absolute change in the feed intake between two consecutive days as $\Delta_{i(j+1)}^{(1)} = FI_{i(j+1)} - FI_{ij}$ for j that satisfies $LD_{i(j+1)} - LD_{ij} = 1$. For instance $\Delta_{i(j+1)}^{(1)} = 0$ means there was no change in feed intake of sow i between the current day and the previous day, while $\Delta_{i(j+1)}^{(1)} < 0$ means that the feed intake consumed by the i th sow in the current day is smaller than the feed intake consumed in the previous day. However, the same amount of change in the feed intake may reflect some stress level for a sow who typically eats a lot and a more serious stress level for a sow that usually has a lower appetite. For this, we define the relative change in the feed intake by $\Delta_{i(j+1)}^{(2)} = (FI_{i(j+1)} - FI_{ij}) / \{(LD_{i(j+1)} - LD_{ij}) \cdot TA_i\}$, where TA_i is the trimmed average of feed intake of i th sow calculated as the average feed intake after removing the lowest 20% and highest 20% of the feed intake measurements taken for the corresponding sow, $FI_{i1}, \dots, FI_{in_i}$. Here TA_i is surrogate for the usual amount of feed intake of the i th sow. Trimmed average is used instead of the common average, to remove outliers of very low feed intakes in first

few lactating days. For example, consider the situation of two sows: sow i that typically consumes 10lb food per day and sow i' that consumes 5lb food per day. A reduction of 5lb in the feed intake over two consecutive days corresponds to $\Delta_{i(j+1)}^{(2)} = -10\%$ for the i th sow and $\Delta_{i'(j+1)}^{(2)} = -20\%$ for the i' th sow. Clearly both sows are stressed (negative value) but the second sow is stressed more, as its absolute value is larger; in view of this we refer to the second response as the *stress index*. Due to the definition of the two types of responses, the data size varies, so for the first response, $\Delta_{i(j+1)}^{(1)}$, we have sample sizes of 233, 350, and 278 for lactation days 4 ($j = 3$), 11 ($j = 10$), and 18 ($j = 17$), respectively, whereas for $\Delta_{i(j+1)}^{(2)}$ the sample sizes are 362, 373, and 336 respectively.

In this analysis we center the attention on the effect of the ambient temperature and humidity on the *1st quartile* of the proxy stress measures and gain more understanding of the food consumption of sows that are most susceptible to heat stress. While the association between the feed intake of lactating sows and the ambient conditions of the farrowing room has been an active research area for some time, accounting for the temperature daily profile has not been considered yet hitherto. Figure 4 displays the temperature and humidity profiles for three days. Preliminary investigation reveals that temperature is negatively correlated with humidity at each time; this phenomenon is caused because the farm uses cool cell panels and fans to control the ambient temperature. Furthermore, it appears that there is strong pointwise correlation between temperature and humidity. In view of these observations we consider the mean summary of the humidity in our analysis. Exploratory analysis of the feed intake behavior of the sows suggest similarities for the sows with parity greater than one (who are at their third pregnancy or higher); thus we use a parity indicator instead of the actual parity of the sow. The parity indicator P_i is defined as one, if the i th sow has parity one and zero otherwise. As emphasized throughout the paper, the existing literature on quantile regression is not suitable to incorporate covariates of different types, as it is the case here.

For the analysis we smooth daily temperature measurements of each sow using uni-

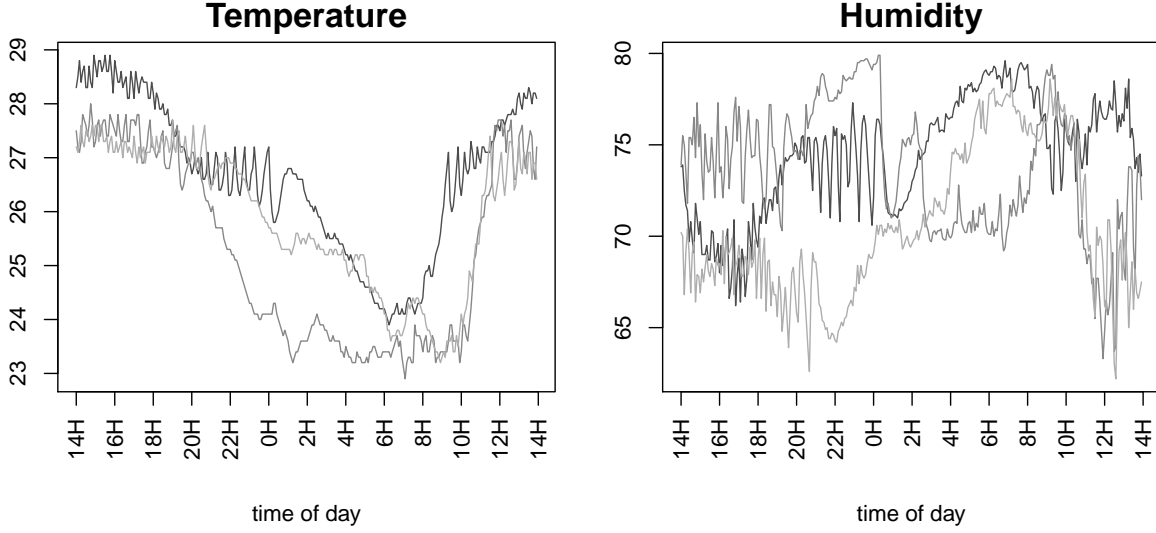


FIG 1. *Temperature ($^{\circ}\text{C}$) and humidity (%) profiles for three randomly selected days. The x-axis begins at 14H (2PM).*

variate smoother with 15 cubic regression bases and quadratic penalty; REML is used to estimate smoothing parameter. The smoothed temperature curve for sow i 's j th repeated measure is denoted by $T_{ij}(t)$, $t \in [0, 24)$. In addition the corresponding daily average humidity is denoted by AH_{ij} . Both temperature and average humidity are centered before being used in the analysis.

In the following we detail estimation procedure. Since the procedure is identical for both responses here we remove superscript in notation and use Δ_{ij} as our response. We first estimate the conditional distribution of Δ_{ij} given temperature $T_{ij}(t)$, average humidity AH_{ij} , parity P_i , and interaction $AH_{ij} \cdot T_{ij}(t)$. Specifically for each of $j = 3, 10$ and 17 we create a set of 100 equi-spaced grid of points between the fifth smallest and fifth largest values of Δ_{ij} 's and denote the grids with $\mathcal{D} = \{d_{\ell} : \ell = 1, \dots, 100\}$. Then we create artificial binary response, $\{\mathbb{1}(\Delta_{ij} < d_{\ell}) : \ell = 1, \dots, 100\}$, and fit the following model for

$$F_{ij}(d_\ell) = E [\mathbb{1}(\Delta_{ij} < d_\ell) | T_{ij}(t), AH_{ij}, P_i]:$$

$$E [\mathbb{1}(\Delta_{ij} < d_\ell) | T_{ij}(t), AH_{ij}, P_i] = g^{-1} \left\{ \beta_0(d_\ell) + \beta_1(d_\ell)P_i + \beta_2(d_\ell)AH_{ij} \right. \\ \left. + \int \beta_3(d_\ell, t)T_{ij}(t)dt + \int \beta_4(d_\ell, t)T_{ij}(t)AH_{ij}dt \right\},$$

where $\beta_0(\cdot)$ is a smooth intercept, $\beta_1(\cdot)$ quantifies the smooth effect of young sows, $\beta_2(\cdot)$ describes the effect of the humidity, and $\beta_3(\cdot, t)$ and $\beta_4(\cdot, t)$ quantify the effect of the temperature at time t as well as the interaction between the temperature at time t and average humidity. We model $\beta_0(\cdot)$ using 20 univariate basis functions, $\beta_1(\cdot)$ and $\beta_2(\cdot)$ using five univariate basis functions, $\beta_3(\cdot, \cdot)$ and $\beta_4(\cdot, \cdot)$ using tensor product of two univariate bases functions (total of 25 functions). Throughout analysis cubic B-spline bases are used and REML is used for estimating smoothing parameters. The estimated conditional distribution, denoted by $\hat{F}_{ij}(d)$, is monotonized by fitting isotonic regression to $\{(d_\ell, \hat{F}_{ij}(d_\ell)) : \ell = 10, \dots, 90\}$; ten smallest and ten largest d_ℓ and the corresponding values of $\hat{F}_{ij}(d_\ell)$ are removed to avoid boundary effects. By abuse of notation we again denote the resulting monotonized estimated distribution with $\hat{F}_{ij}(d)$. Finally we obtain estimated quantiles at $\tau = 0.25$ level by inverting the estimated distribution function $\hat{F}_{ij}(d)$, namely $\hat{Q}(\tau = 0.25 | T_{ij}(t), AH_{ij}, P_i) = \inf\{d : \hat{F}_{ij}(d) > 0.25\}$.

To study and interpret the effect of each covariate we predict quantiles at combinations of different values of covariates and investigate the predicted values against the covariates. For each of three lactation days ($j = 3, 10, 17$) we consider three values of average humidity (first quartile, median, and third quartile) and the two levels of parity (0 for older sows and 1 for younger sows). For the functional covariate $T_{ij}(t)$ we create seven smooth temperature curves given in Figure 2 by first obtaining pointwise quantiles of temperature and then smoothing them for each of quantiles levels $\eta = 0.2, 0.3, \dots$, and 0.8. In summary for each of three lactation days we make prediction of distribution and the first quartile of two responses for 42 ($3 \times 2 \times 7$) different combinations of covariates values. To avoid extrapolation we ascertain that there are reasonably many observed measurements

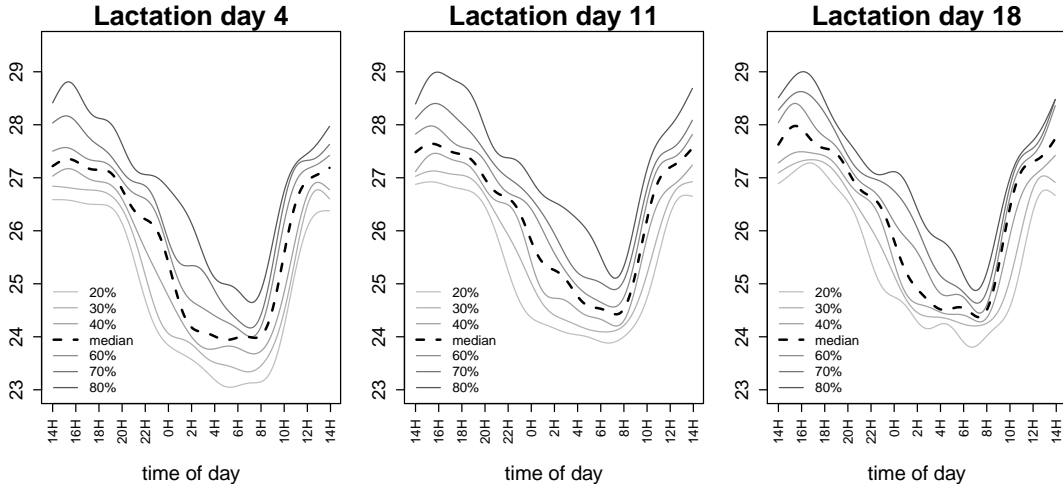
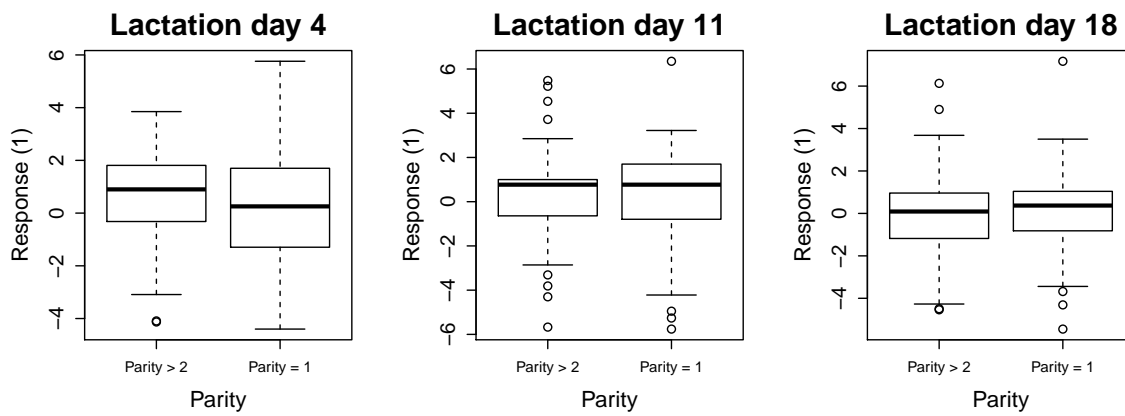


FIG 2. Temperature curves with which prediction of quantiles is made. Dashed black line is pointwise average of temperature curves and solid lines are pointwise quantiles; all curves are smoothed.

at each of the combinations and bottom 25% of the responses are not dominantly from one of the parity group; see distribution of each response by the parity in Figure 3.

The resulting predicted quantiles are shown in Figure 4. Here we focus our discussion on predicted quantile of $\Delta_{i(j+1)}^{(2)}$ at quantile level $\tau = 0.25$ for lactation day 4 ($j = 3$) - the first plot of the second row in Figure 4. The results suggest that the feed intake of older sows (parity indicator equal to zero) are less affected by high temperatures than that of younger sows; this finding is in agreement with Bloemhof et al. (2013), who also found that younger sows (parity equal to one) are more sensitive to ambient changes than sows with higher parity. We also observe that humidity plays an important role in the effect of temperature on feed intake change. Similar to a previous study (Bergsma and Hermes, 2012) our results also imply that high humidity is related to a negative impact of high temperature on feed intake while low humidity alleviates it. On the contrary, when temperature is low, high humidity leads to better feed intake than low humidity. For instance on lactation day 4 ($j = 3$) regardless of the parity group, when temperature increases the predicted first quartile of $\Delta_{i(j+1)}^{(2)}$ increases for low humidity (solid line) whereas it decreases for high humidity (dotted line). The result seems to suggest to keep low humidity levels in

Response (1) by Parity



Response (2) by Parity

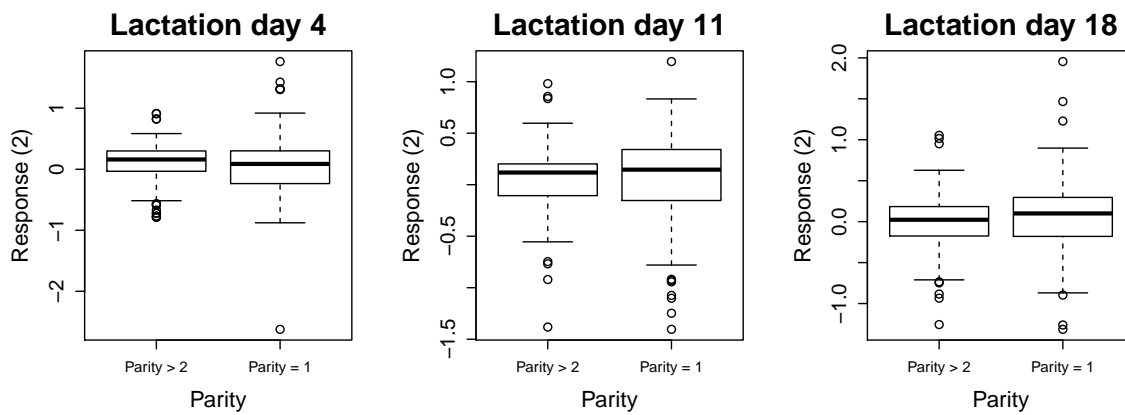


FIG 3. Distribution of responses by Parity

order to maintain healthy feed intake behavior, when ambient temperature is above 60th percentile; high humidity levels are desirable for cooler ambient temperature.

The results corresponding to other lactation days can also be interpreted similarly. While the effects of covariates on feed intake are less apparent toward the end of lactation period, we still observe similar pattern across all three lactation days. For lactation day 11 ($j = 10$) we observe that when temperature is above 40th percentile the predicted first quartile starts increasing with low humidity while it continues decreasing with high humidity. Similarly for lactation day 18 ($j = 17$) when temperature is above 60th percentile the predicted first quartile increases with low humidity while it decreases with high humidity. The effect of temperature on feed intake seems less obvious for lactation days 11 and 18 than for day 4; while the effect may be due to lactation day it may also be a result of other factors, such as more fluctuation and variability in temperature curves on day 4 than other two days (see Figure 2). Overall we conclude that high humidity and temperature affect sows' feed intake behavior negatively and young sows (parity one) are more sensitive to heat stress than older sows (higher parity), especially in the beginning of lactation period.

5. Discussion

In this paper we proposed a novel framework for a comprehensive study of covariates of mixed types on the conditional distribution of the response. Extensive simulation study showed very good prediction performance in terms of estimating quantiles of various levels. Additionally, the modeling approach leads to computationally efficient estimation algorithm. The proposed method is flexible and easy to implement using existent software, `pffr` (Scheipl et al., 2015).

This modeling framework opens up a couple of future research directions. A first research avenue is to develop significance tests of null covariate effect. Testing for the null effect of a covariate on the conditional distribution of the response is equivalent to testing

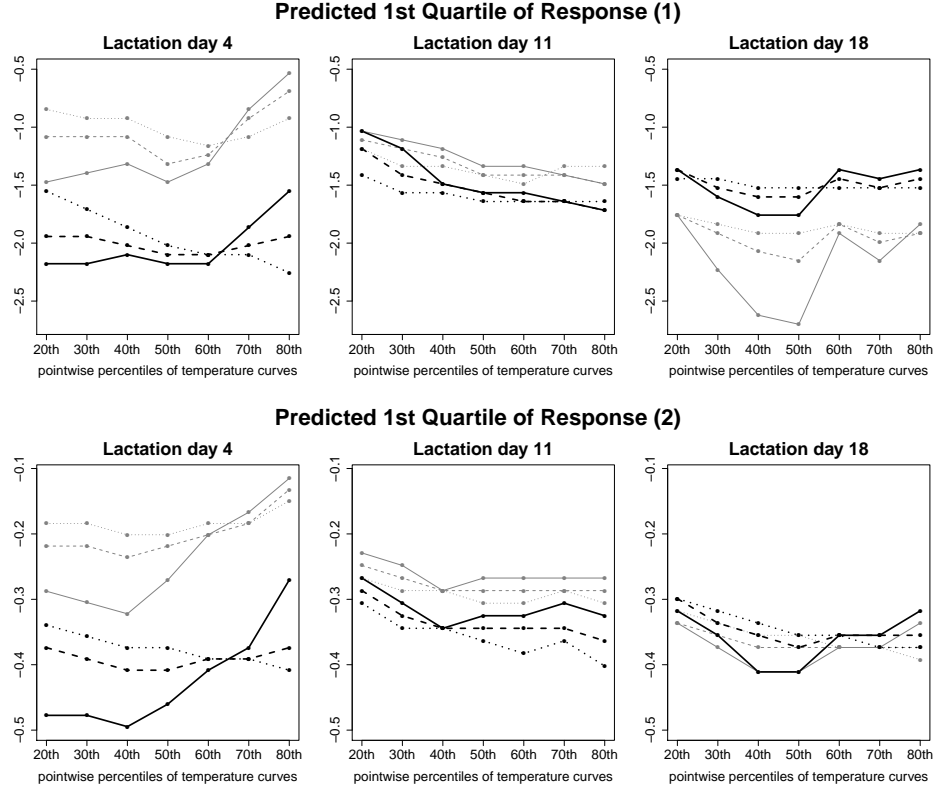


FIG 4. Displayed are the predicted quantiles of $\Delta_{i(j+1)}^{(1)}$ and $\Delta_{i(j+1)}^{(2)}$ for different parities, average humidity, and temperature levels. In each of all six panels, black thick lines correspond to the young sows ($P_i = 1$) and grey thin lines correspond to the old sows ($P_i = 0$). Line types indicate different average humidity levels; solid, dashed, and dotted correspond to low, medium, and high average humidity levels (given by the first quartile, median, and the third quartiles of AH_{ij}), respectively. The seven grids in x-axis of each panel correspond to the 7 temperature curves given in the respective panel of Figure 2.

that the corresponding regression coefficient function is equal to zero in the associated function-on-function mean regression model. Such significance tests have been studied when the functional response is continuous (Shen and Faraway, 2004; Zhang et al., 2007); however their study for binary-valued functional responses is an open problem in functional data literature. One possible alternative is to construct confidence bands for the corresponding coefficient function, say using bootstrap, and examine whether the confidence band includes zero for its entire domain. Another research avenue is to do variable selection in the setting where there are many scalar covariates and functional covariates. Many current applications collect data with increasing number of mixed covariates and selecting the ones that have an effect on the conditional distribution of the response is very important. This problem is an active research area in functional mean regression where the response is normal (Gertheiss et al., 2013; Chen et al., 2015). The proposed modeling framework has the potential to facilitate studying such problem.

Supplementary Material

Section S1 provides additional simulation settings and results for the cases of having either a single scalar covariate or a single functional covariate. Section S2 presents additional data analysis done using the proposed method on the bike sharing dataset (Fanaee-T and Gama, 2013; Lichman, 2013). Lastly the R function implementing the proposed method is available at http://www4.ncsu.edu/~spark13/software/QRFD_Rcode.zip/.

Acknowledgement

Staicu's research was supported partially by National Science Foundation DMS 0454942 and National Institutes of Health grants R01 NS085211 and R01 MH086633. The data used originated from work supported in part by the North Carolina Agricultural Foundation, Raleigh, NC.

References

- Barlow, R. E., Bartholomew, D. J., Bremner, J., and Brunk, H. D. (1972). *Statistical inference under order restrictions: the theory and application of isotonic regression*. Wiley New York.
- Bergsma, R. and Hermes, S. (2012). Exploring breeding opportunities for reduced thermal sensitivity of feed intake in the lactating sow. *Journal of animal science*, 90(1):85–98.
- Black, J., Mullan, B., Lorschy, M., and Giles, L. (1993). Lactation in the sow during heat stress. *Livestock production science*, 35(1):153–170.
- Bloemhof, S., Mathur, P., Knol, E., and Van der Waaij, E. (2013). Effect of daily environmental temperature on farrowing rate and total born in dam line sows. *Journal of animal science*, 91(6):2667–2679.
- Bondell, H., Reich, B., and Wang, H. (2010). Noncrossing quantile regression curve estimation. *Biometrika*, 97(4):825–838.
- Cardot, H., Crambes, C., and Sarda, P. (2005). Quantile regression when the covariates are functions. *Nonparametric Statistics*, 17(7):841–856.
- Cardot, H., Ferraty, F., and Sarda, P. (1999). Functional linear model. *Statistics & Probability Letters*, 45(11 – 22).
- Chen, K. and Müller, H.-G. (2012). Conditional quantile analysis when covariates are functions, with application to growth data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(1):67–89.
- Chen, Y., Goldsmith, J., and Ogden, T. (2015). Variable selection in function-on-scalar regression.
- Chernozhukov, V., Fernandez-Val, I., and Galichon, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika*, page asp030.
- Crainiceanu, Philip Reiss, Jeff Goldsmith, Lei Huang, Lan Huo, and Fabian Scheipl (2013).

- refund: Regression with Functional Data*. R package version 0.1-9.
- Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, pages 89–102.
- Fanaee-T, H. and Gama, J. (2013). Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, pages 1–15.
- Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic "el niño" phenomenon. *Sankhyā: The Indian Journal of Statistics*, pages 378–398.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice*. Springer Science & Business Media.
- Ferraty, F. and Vieu, P. (2009). Additive prediction and boosting for functional data. *Computational Statistics & Data Analysis*, 53(4):1400–1413.
- Gertheiss, J., Maity, A., and Staicu, A.-M. (2013). Variable selection in generalized functional linear models. *Stat*, 2(1):86–101.
- Goldsmith, J., Bobb, J., Crainiceanu, C. M., Caffo, B., and Reich, D. (2011). Penalized functional regression. *Journal of Computational and Graphical Statistics*, 20(4).
- Goldsmith, J., Bobb, J., Crainiceanu, C. M., Caffo, B., and Reich, D. (2012). Penalized functional regression. *Journal of Computational and Graphical Statistics*.
- Ivanescu, A. E., Staicu, A.-M., Scheipl, F., and Greven, S. (2015). Penalized function-on-function regression. *Computational Statistics*, pages 1–30.
- James, G. M. (2002). Generalized linear models with functional predictors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):411–432.
- Johnston, L., Ellis, M., Libal, G., Mayrose, V., and Weldon, W. (1999). Effect of room temperature and dietary amino acid concentration on performance of lactating sows. ncr-89 committee on swine management. *Journal of animal science*, 77(7):1638–1644.
- Kato, K. et al. (2012). Estimation in functional linear quantile regression. *The Annals of Statistics*, 40(6):3108–3136.

- Koenker, R. (2005). *Quantile regression*. Number 38. Cambridge university press.
- Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.
- Lichman, M. (2013). UCI machine learning repository.
- Lu, Y., Du, J., and Sun, Z. (2014). Functional partially linear quantile regression model. *Metrika*, 77(2):317–332.
- Marx, B. D. and Eilers, P. H. (2005). Multidimensional penalized signal regression. *Technometrics*, 47(1):13–22.
- MATLAB (2010). *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts.
- McLean, M., Hooker, G., Staicu, A.-M., Scheipl, F., and Ruppert, R. (2014). Functional generalized additive models. *Journal of Computational and Graphical Statistics*, 23:249–269.
- Melillo, J. M., Richmond, T. T., and Yohe, G. (2014). Climate change impacts in the united states. *Third National Climate Assessment*.
- Nelson, G. C., Rosegrant, M. W., Koo, J., Robertson, R., Sulser, T., Zhu, T., Ringler, C., Msangi, S., Palazzo, A., Batka, M., et al. (2009). *Climate change: Impact on agriculture and costs of adaptation*, volume 21. Intl Food Policy Res Inst.
- Ng, P. and Maechler, M. (2007). A fast and efficient implementation of qualitatively constrained quantile smoothing splines. *Statistical Modelling*, 7(4):315–328.
- Quiniou, N. and Noblet, J. (1999). Influence of high ambient temperatures on performance of multiparous lactating sows. *JOURNAL OF ANIMAL SCIENCE-MENASHA THEN ALBANY THEN CHAMPAIGN ILLINOIS-*, 77:2124–2134.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsay, J. O. (2006). *Functional data analysis*. Wiley Online Library.
- Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis: methods and case studies*, volume 77. Citeseer.

- Renaudeau, D. and Noblet, J. (2001). Effects of exposure to high ambient temperature and dietary protein level on sow milk production and performance of piglets. *Journal of animal science*, 79(6):1540–1548.
- Renaudeau, D., Quiniou, N., and Noblet, J. (2001). Effects of exposure to high ambient temperature and dietary protein level on performance of multiparous lactating sows. *Journal of Animal Science*, 79(5):1240–1249.
- Rosero, D. S., Boyd, R. D., McCulley, M., Odle, J., and van Heugten, E. (2016). Essential fatty acid supplementation during lactation is required to maximize the subsequent reproductive performance of the modern sow. *Animal Reproduction Science*, 168:151–163.
- Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, 11(4).
- Scheipl, F., Gertheiss, J., and Greven, S. (2015). Generalized functional additive mixed models. *arXiv preprint arXiv:1506.05384*.
- Shen, Q. and Faraway, J. (2004). An f test for linear models with functional responses. *Statistica Sinica*, pages 1239–1257.
- St-Pierre, N., Cobanov, B., and Schnitkey, G. (2003). Economic losses from heat stress by us livestock industries. *Journal of dairy science*, 86:E52–E77.
- Tang, Q. and Cheng, L. (2014). Partial functional linear quantile regression. *Science China Mathematics*, 57(12):2589–2608.
- Wood, S. (2006a). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.
- Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):95–114.
- Wood, S. N. (2006b). Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, 62(4):1025–1036.
- Xiao, L., Li, Y., and Ruppert, D. (2013). Fast bivariate p-splines: the sandwich smoother.

- Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(3):577–599.
- Xiao, L., Zipunnikov, V., Ruppert, D., and Crainiceanu, C. (2016). Fast covariance estimation for high-dimensional functional data. *Statistics and Computing*, 26(1-2):409–421.
- Yao, F., Müller, H.-G., and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470):577–590.
- Yu, D., Kong, L., and Mizera, I. (2015). Partial functional linear quantile regression for neuroimaging data analysis. *arXiv preprint arXiv:1511.00632*.
- Zhang, J.-T., Chen, J., et al. (2007). Statistical inferences for functional data. *The Annals of Statistics*, 35(3):1052–1079.